

# Technical Note

## Counting FW 6.30



**BOSCH**  
Invented for life

April 21st, 2016

### Counting with Essential Video Analytics 6.30 / Intelligent Video Analytics 6.30

#### Summary

This technical note describes how to count with Essential Video Analytics 6.30 and Intelligent Video Analytics 6.30, what the best perspective is, how to set it up, how to get the data from the camera, and what the best practices are.

#### Applications

- Determine conversion rates in shops
- Alert on queue lengths before cashiers
- Count vehicles for traffic monitoring

#### What can be counted?

Via the GUI, two dedicated counting tasks are available with FW 6.30:

- Counter:
  - Count all objects crossing a specified line
  - The line can consist of up to 16 edge points or 15 segments
  - An alarm can be raised if the counter reaches a maximal value and should be reset via RCP+
- Occupancy:
  - Count all objects currently in a specified field
  - The field can have up to 16 vertices and edges
  - An alarm can be raised if a predefined threshold is reached

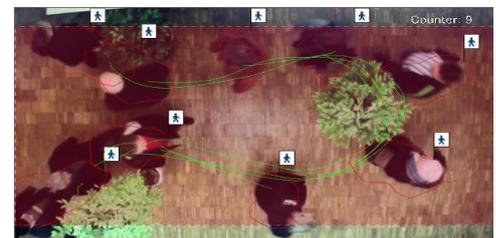
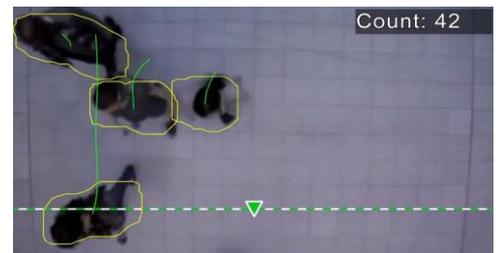
Both counter can be applied to arbitrary objects and restricted according to the objects properties, e.g. by the object class to separate vehicles from bikes and pedestrians.

Via the VCA task script language, counter can be applied to any event (like entering a field) defined there and any object state possible. It is also possible to define a virtual room there via up to three entry and exit lines. See the tech note on VCA task script language for further details and examples.

People counting is supported via a special 3D people tracking mode for accurate results even in well-populated scenes.

#### Limitations

- General objects need to be well separated for the counter to work accurately. If objects occlude each other, they will be combined into a single objects, skewing the results.
- Shadows, reflections, moving background like vegetation and light changes will decrease performance.
- Use 3D people tracking to accurately separate people even in well-populated scenes. 3D people tracking requires accurate calibration. It assumes that every moving object is a person or group of persons, including shopping carts, and will thus generate extra counts for non-person objects.



Counting people crossing a line (top) or currently inside a field (bottom). 3D people tracking mode allows to accurately separate close and slightly occluded people even in well-populated scenes

## Perspective: Why bird's eye view is preferred



### Front view:

- ✗ A lot of people are occluded by other people
- ✗ Very hard to detect & separate people in the rear
- ✗ Position in the room cannot be determined
- ➔ Not usable for counting at all



### Slight tilting:

- ✗ People are mostly occluded
- ✗ Very hard to detect & separate people in the rear
- ✗ Position in the room very ambiguous
- ➔ Not usable for counting at all



### More tilting:

- ✗ People are partly occluded
- ✓ Position in the room can be determined
- ➔ Human can count here manually
- ➔ Automated computer counting limited by people density



### Bird's eye view:

- ✓ All people can be seen and separated easily
- ✓ Position in the room easy to determine
- ➔ Max. quality of the counting results

### People counting best practices: Planning camera field of view

Checklist	Comments
Use birds eye view if possible	See the following chapter for explanation. Note that performance will decrease the more the camera view differs from birds eye view (90° tilt angle).
Camera height ~3-6 meter (also depending on camera lens)	Less height and the people walk through the observed area too fast to be detected and counted properly. More height and the people become too small, reducing accuracy.
Uniform illumination	Changing illumination triggers false foreground especially at edges up to a global change where nothing is processed anymore
Ambient light	Shadows of persons move with them and might be counted falsely. No shadows are visible in ambient light.
Indoor only	To keep illumination stable, uniform and ambient.
No reflecting surfaces	Reflections of persons move with them and might be counted falsely.
No moving background, e.g. doors, glass windows, escalators, shopping card / basket collecting points,...	Moving background always yields false foreground segmentation. Take these areas out of the sensitive area or completely out of the image. Besides generating false people themselves, they can also delay the detection of objects entering nearby. E.g. a track started by a person moving through a door might stay at the door, delaying the detection of the person.
No baggage, shopping carts, cars,...	All non-person objects of size similar to or larger than persons will be counted as persons.
Max number of persons in scene at once for Intelligent Video Analytics: 10 (CPP4) / 20 (CPP6/7), for Essential Video Analytics: 10 (CPP7)	If more persons are in the camera scene at once, real-time performance can no longer be guaranteed. This means that single frames might be dropped, though object tracks will be continued. This may lower overall performance.
Use straight passageways where possible	Avoid scenes where people enter, directly turn around a corner and immediately leave the camera image before the algorithm had a change to detect and track them robustly. Avoid scenes where people loiter and thus cross the counting line several times.

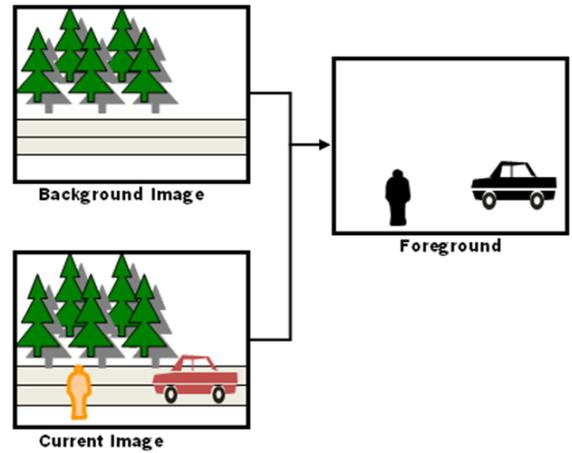
**Technology: Background Subtraction & 3D Models**

**Background Subtraction**

Essential Video Analytics and Intelligent Video Analytics are currently based on a technique named background subtraction. There, a background image is learned (and updated over time) and any difference to the background is classified as moving foreground.

Advanced processing allows for illumination invariance up to some point.

Unfortunately, unwanted “disturbers” like trees, shadows, reflections, opening doors and more are also changes in comparison to the background...



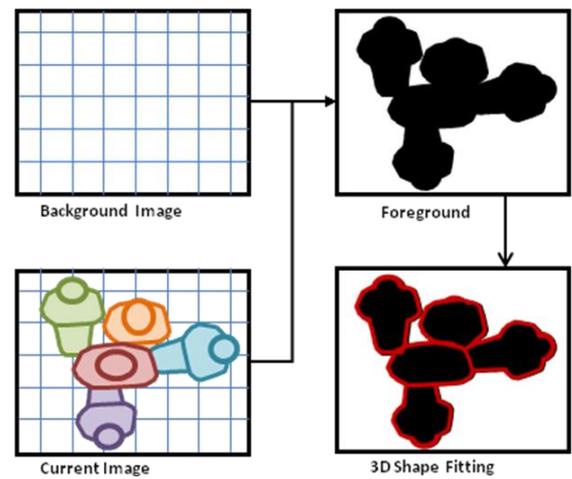
**3D people tracking: Object Separation via person models**

Background subtraction of a group of objects results in a single foreground blob. Further separation of the objects is necessary for counting.

For this, all objects are interpreted as persons or groups of persons. 3D person models are calculated according to camera calibration and fitted to the extracted foreground.

Good separation results (and thus also counting accuracy) depend on a good calibration.

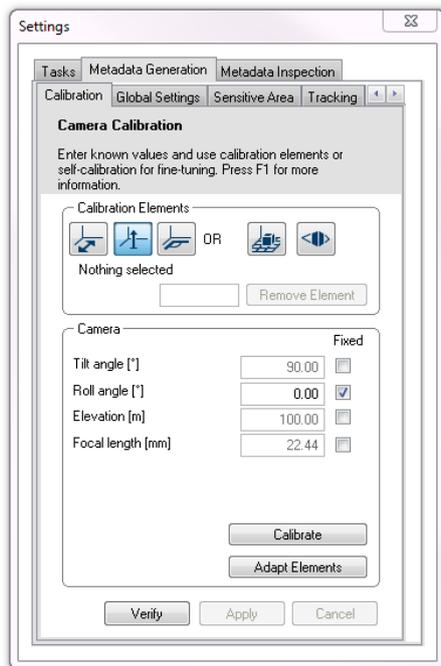
To initialize and separate objects, enough foreground has to be generated to support the fit of the 3D person model to the foreground. This may delay the detection of persons.



## Calibrating the camera

Calibration is needed if 3D people tracking is to be used for accurate results for people counting, or if 3D tracking is needed to automatically generate object classes to filter the counts.

To calibrate the camera for a bird's eye view perspective, open the Calibration Dialog and fix all parameters. Tune the parameters manually using the description below. The grid visualization can be used for verification by aligning a cube with vertical structures, or with tiles on the ground. For final verification, set the calibration, enable 3D (people) tracking mode (see also Enabling People Counting) and check the fit of the resulting 3D person model shapes on people walking through the observed scene (see also Example Configurations). Note that especially the 3D person model size will later influence people counting accuracy and should be used to tune a trade-off between separation of people piggy-backing and adding false positives on baggage.



Parameter	Comment
Tilt Angle	90° = Birds eye view 0° = Fully frontal view Vertical line alignment    <-> \/
Roll Angle	Left/right camera rotation Vertical line alignment \ \ <->    <-> //
Elevation	Please measure camera height Can be used for 3D person model size fitting
Focal length	See camera lens Can be used for 3D person model size fitting

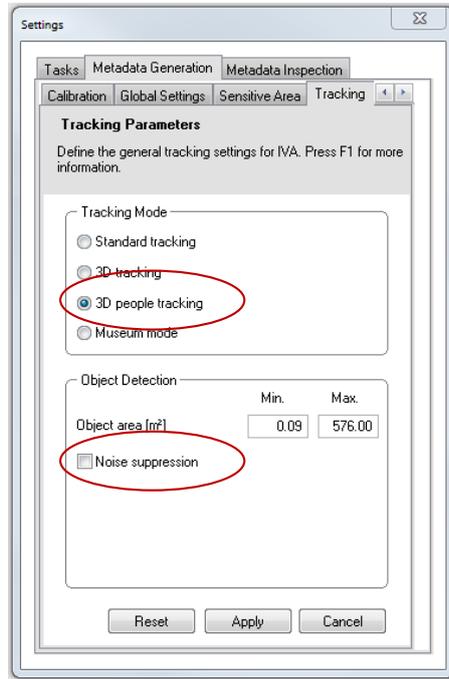
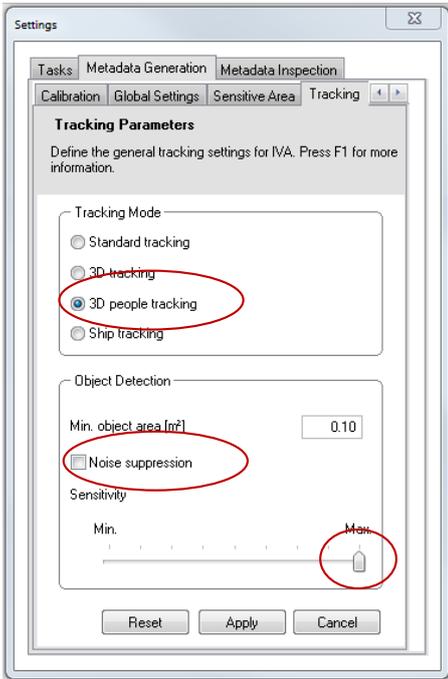
Please note that certain objectives induce a distortion, which cannot be corrected and makes calibration more difficult. The tile pattern in the image on the right is a good example for that, for a perfect calibration all lines should be straight. In these cases, approximate as good as you can, and focus on the area with the counting line.



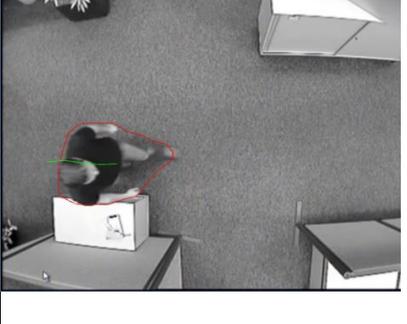
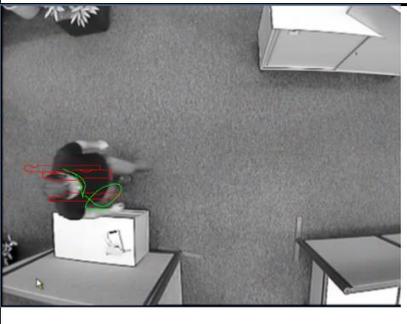
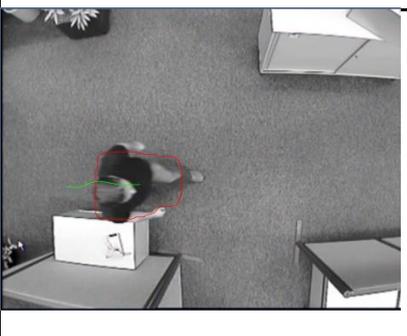
### Enabling people counting

Set the following parameters:

Parameter	Value	Effect & Comments
3D people counting	On	Use 3D person model, interpret everything as persons, separate all persons. Will only be available if camera is calibrated.
Noise suppression	Off	Noise suppression can delay detection and will not work well in crowds.
Sensitivity (Intelligent Video Analytics on CPP6/7 only)	Max	Initialize new persons immediately instead of validating them over time. Needed especially for bird's eye view as the people are only in the FOV for a short time.



Example Configurations

Example 3D Shapes	Configuration Status	What To Do
	<ul style="list-style-type: none"> <li>✗ 3D people tracking not enabled</li> </ul>	<ul style="list-style-type: none"> <li>➔ Enable 3D people tracking</li> </ul>
	<ul style="list-style-type: none"> <li>✓ 3D people tracking is enabled and the 3D person model is visible</li> <li>✗ 3D person models too large</li> <li>✗ 3D person models indicates frontal view in contrast to the images BEV</li> </ul>	<ul style="list-style-type: none"> <li>➔ Set Tilt to ~90</li> <li>➔ Adapt elevation / focal length</li> </ul>
	<ul style="list-style-type: none"> <li>✓ 3D people tracking is enabled and the 3D person model is visible</li> <li>✗ 3D person models too small, several 3D person models on one person</li> <li>✗ 3D person models indicates frontal view in contrast to the images BEV</li> </ul>	<ul style="list-style-type: none"> <li>➔ Set Tilt to ~90</li> <li>➔ Adapt elevation / focal length</li> </ul>
	<ul style="list-style-type: none"> <li>✓ 3D people tracking enabled and the 3D person model is visible</li> <li>✗ 3D person models too small, several 3D person models on one person</li> <li>✓ 3D person models indicate BEV calibration which fits the image</li> </ul>	<ul style="list-style-type: none"> <li>➔ Adapt elevation / focal length</li> </ul>
	<ul style="list-style-type: none"> <li>✓ 3D people tracking enabled and the 3D person model is visible</li> <li>✓ 3D person models fits the person very well</li> <li>✓ 3D person models indicate BEV calibration which fits the image</li> </ul>	

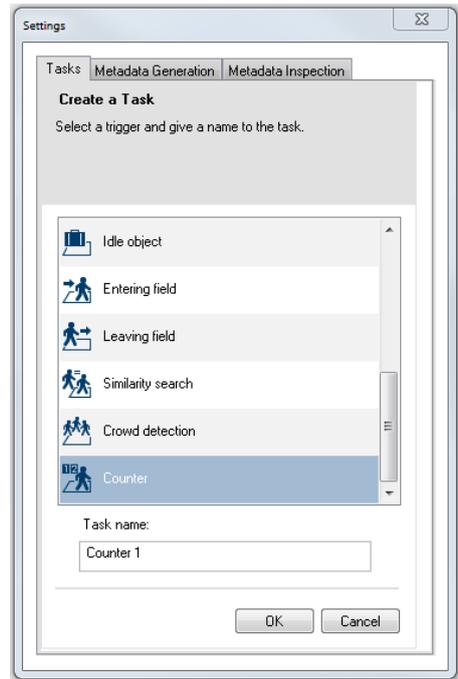
### Setup of Counting Lines / Fields

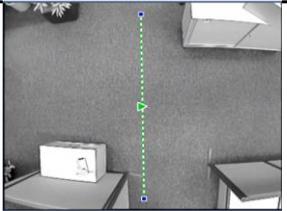
To set up a people counter, create a counter task or an occupancy task. Counting lines or fields can be defined via the corresponding wizard. Note that the lines and fields will be shown in the camera image and can be changed there.

Set the counting lines / fields up thus that objects can be well detected both before and after crossing the line / field border. This ensures observation of the person crossing of the line / field border with high quality needed for high counting accuracy.

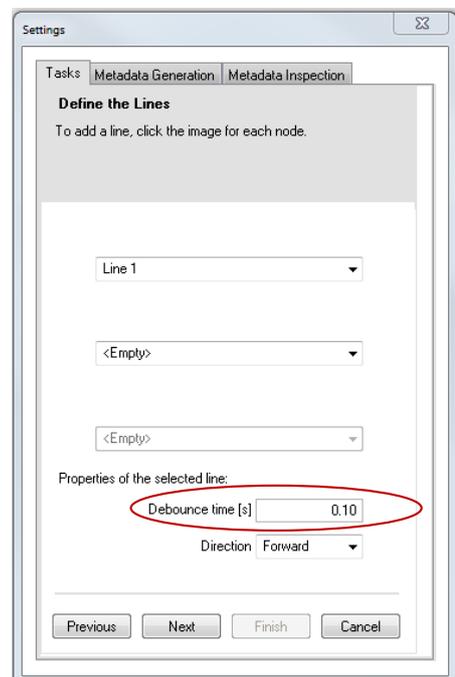
Make sure the counting lines / fields are large enough that people can't slip by. Note that the center of the object shape is used to determine line / field border crossing.

Note that all visualizations in the camera image can be selected and dragged around.



Examples	Comments
	<ul style="list-style-type: none"> <li>✗ Line too close to border. Objects entering will not have generated enough foreground to be detected.</li> <li>✗ Not suitable for counting.</li> </ul>
	<ul style="list-style-type: none"> <li>✓ Enough room before and after the counting line. Objects will be detected with high accuracy before crossing the counting line and the crossing itself will be observed with high quality.</li> <li>✓ Well set up for counting.</li> </ul>

Set the debounce time of line or field to 0.1. Debounce time is useful to avoid alarm flickering in case of objects moving along a detection line or field border. In order to do this, it enforces a certain duration of the detected object before and after crossing the line. In a bird's eye view (BEV), the camera is very close to the persons and thus the persons remain only a short time within the monitored area. Therefore the duration needed for debounce is often not available, and thus a large debounce time cuts down the number of correctly counted people. A short debounce time, however, is needed to not count people several times which stop at the counting line.

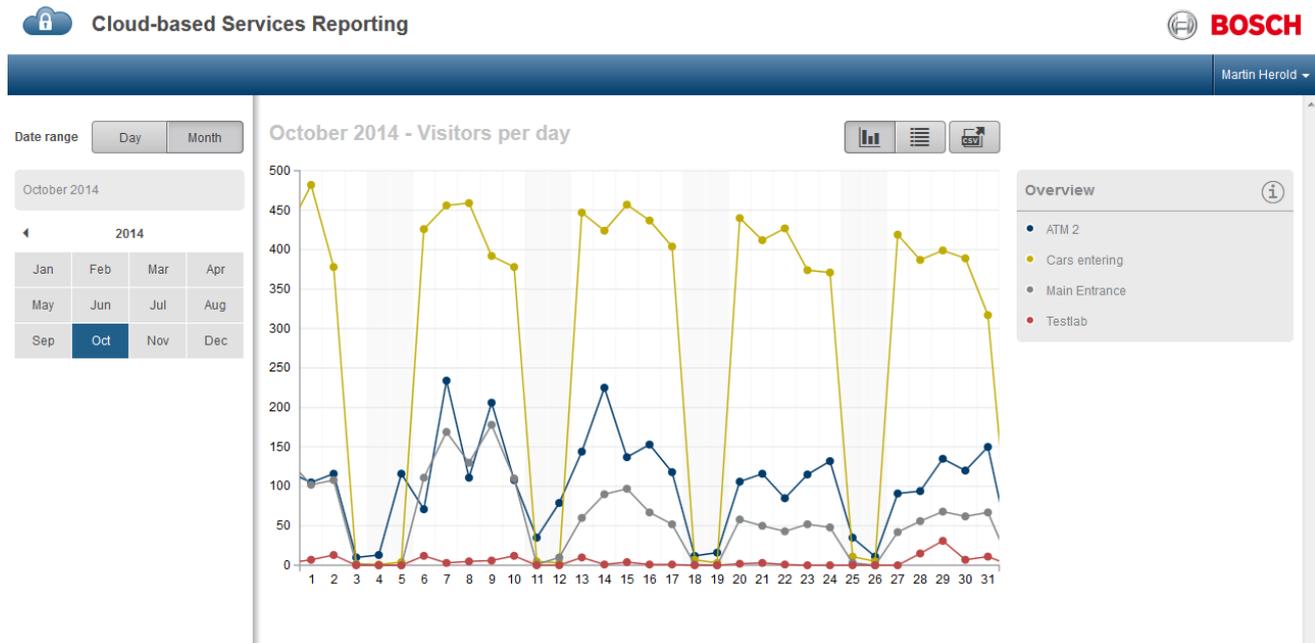


Do not worry if the checkbox of the final counter task is not set and can't be activated - this means only that the counter will generate no alarm. If an alarm is needed, you can add it in the counter task wizard. Otherwise, just ignore the checkbox as the counting will work nonetheless.



## Accessing the counter data: CBS

Bosch Cloud Based Security and Services (CBS) offers a monthly and daily customer traffic report:



## Accessing the counter data: RCP+, VCA metadata

Remote Control Protocol plus (RCP+) is a remote control protocol including the ability to send read and write commands to a device, and to subscribe for messages. Defined counters can be read out and reset via RCP+ commands. It is best practice to reset the counter values regularly so they will not generate an overflow of the counter values.

The VCA metadata stream including all detected and tracked objects as well as the counter data is encoded in the proprietary Video Content Description (VCD) Format and fully accessible via Remote Control Protocol plus (RCP+) and Real Time Streaming Protocol (RTSP). Via ONVIF and in the ONVIF metadata description format, a reduced set of the VCA metadata including the events and object bounding boxes is available. As the VCA metadata contains the full information about all detected and tracked objects, this can be used to define further counters or generate heatmaps in any client able to read the VCA metadata.

Contact the Integration Partner Program (<http://ipp.boschsecurity.com>) for a full documentation or exemplary source code.